JOURNAL OF COMPUTER AND SYSTEM SCIENCES: 1, 235-240 (1967)

First-Order Necessary Conditions for Generalized Optimization Problems

T. GUINN

Mathematics Department, Michigan State University, East Lansing, Michigan 48823

and

Douglas Aircraft Co, Santa Monica, California

Received September 6, 1966

1. INTRODUCTION

In Reference [*1*] the question of existence of solutions was considered for a generalized optimization problem. In this paper, necessary conditions for a generalized solution to be optimal will be derived. Since the method of proof will closely follow that of Hestenes [*2*] for optimal control problems, the generalized problem will be posed somewhat differently than in [*1*]. The two can easily be shown to be equivalent.

Let T be an open interval of the real line, V an open set in E^n, B an open set in E^p. Let $\mathscr{F} = \{f(t, x)\}$ be a set of $(n + 1)$-dimensional vector-valued functions defined on $T \times V$. Let $T^0(b)$ and $T^1(b)$ be two functions defined for b in B with range in T. Let $X^0(b)$ and $X^1(b)$ be two functions defined for b in B with range in V. We will call

$$x\colon x(t),\ f(t, x),\ b$$

a solution if $x(t)$ is an absolutely continuous function in V for t in T, $f(t, x)$ is in $\mathscr{F}$, b is in B and

$$\dot{x}^j(t) = f^j(t, x(t)) \qquad (t^0 \leqslant t \leqslant t^1;\ j = 1,..., n), \tag{1.1}$$

$$t^s = T^s(b),\ x(t^s) = X^s(b) \qquad (s = 0, 1). \tag{1.2}$$

Among all such solutions we seek one for which

$$I_0(x) = \int_{t^0}^{t^1} f^0(t, x(t))\, dt = \min. \tag{1.3}$$

This formation corresponds to a generalization of the control problem of Lagrange, whereas the formulation in [*1*] corresponds to the control problem of Mayer.

2. NECESSARY CONDITIONS

Assume there exists a function $k(t)$, integrable on T, such that, for every f in $\mathscr{F}$,

$f(t, x)$ and f_{x^i} $(i = 1,..., n)$ is Lebesgue-integrable in t on T for fixed x in V, (2.1a)

$$|f(t, x) - f(t, y)| \leqslant k(t)|x - y| \qquad (t, x), (t, y) \in T \times V, \tag{2.1b}$$

$$|f(t, x)| \leqslant k(t) \qquad (t, x) \in T \times V. \tag{2.1c}$$

We omit the qualification "almost everywhere" here and elsewhere, since it is clear where it applies.

Assume that $\mathscr{F}$ has the property that, if f_1 and f_2 belong to $\mathscr{F}$ and T' is any subinterval of T, the function

$$C(T')f_1 + C(T - T')f_2$$

also belongs to $\mathscr{F}$, where C is the characteristic function. We also assume that $T^s(b)$ and $X^s(b)$ have continuous partial derivatives with respect to b on B.

THEOREM 2.1. *Let x_0 afford a strong minimum to J_0 in the class of solutions. Then there exist multipliers $\lambda_0, p_1(t),..., p_n(t)$ such that, if we set*

$$H(t, x, f, p) = p_i f^i - \lambda_0 f^0: \tag{2.2}$$

(1) *The multipliers $\lambda_0, p_i(t)$ do not vanish simultaneously and $\lambda_0 \geqslant 0$;*

(2) *On x_0 the multiplier $p_i(t)$ satisfy*

$$\dot{p}_i(t) = -H_{x^i}; \tag{2.3}$$

(3) *the inequality*

$$H(t, x_0(t), f(t, x_0(t)), p(t)) \leqslant H(t, x(t), f_0(t, x_0(t)), p(t)) \tag{2.4}$$

holds for all f in $\mathscr{F}$;

(4) *the transversality condition*

$$[-H\, dT^s + p_i(t^s)\, dX^{is}]_{s=0}^{s=1} = 0 \tag{2.5}$$

is an identity in db_σ on x_0.

We will prove the theorem for the case where the functions $f(t, x)$ are piecewise continuous in t and where conditions (1.2) are fixed and do not depend on b. Extension to include conditions (1.2) is just as in [2], and to where $f(t, x)$ is integrable in t is as in [3].

Define

$$A_j^i(t) = \frac{\partial}{\partial x^j} f_0^i(t, x_0(t)) \tag{2.6}$$

$$(i, j = 1,..., n).$$

$$B_j(t) = \frac{\partial}{\partial x^j} f_0^0(t, x_0(t)) \tag{2.7}$$

By hypothesis, A_j^i, B_j are integrable on $t^0 \leqslant t \leqslant t^1$. Let $q(t)$ be solutions on $t^0 \leqslant t \leqslant t^1$ of

$$\dot{q}_i + q_j A_i^j(t) = B_i(t), \; q_i(t^0) = 0. \tag{2.8}$$

Then $q(t)$ is an absolutely continuous function which satisfies (2.8). If x is a solution,

$$\frac{d}{dt}(q_i x^i) = B_i x^i + q_i(f^i - A_j{}^i x^j).$$

Set

$$F^0 = f^0 - B_j x^j - q_i(f^i - A_j{}^i x^j), \tag{2.9}$$

$$G^0 = q^i(t^1)X^{i1}. \tag{2.10}$$

By assumption, F^0 is integrable,

$$\begin{aligned} I_0 &= [q_i x^i]_{t}^{t_0^1} + \int_{t^0}^{t^1} \left[f^0 + \frac{d}{dt}(q_i x^i) \right] dt \\ &= G^0 + \int_{t^0}^{t^1} F^0(t, x(t))\, dt \\ &\equiv J_0 \end{aligned}$$

for all solutions x. Hence,

$$J^0(x) - J_0(x_0) = I_0(x) - I_0(x_0).$$

Let $P_{ij}(t)$ be solutions of

$$\dot{P}_{ij} = -P_{ik}A_j{}^k,\ P_{ij}(t^1) = \delta_{ij}\,. \tag{2.11}$$

Then, for any solution x,

$$\frac{d}{dt}(P_{ij}x^j) = P_{ij}(f^j - A_k{}^j x^k).$$

Now set

$$F^i = P_{ij}(f^j - A_k{}^j x^k), \tag{2.12}$$

$$G^i = -[P_{ij}(t^s)\, X^{is}]_{s=0}^{s=1}, \tag{2.13}$$

$$J_i = G^i + \int_{t^0}^{t^1} F^i\, dt.$$

For every solution x,

$$\begin{aligned} J_i &= G^i + \int_{t^0}^{t^1} F^i(t, x(t))\, dt \\ &= [P_{ij}(t^s)(x^j(t^s) - X^{js})]_{s=0}^{s=1}. \end{aligned}$$

Therefore, if $x^j(t^0) = X^{j0}$, $J_i = 0$ if and only if $x^j(t^1) = X^{j1}$. We note that F^ρ was constructed so that, along x_0,

$$\partial F^\rho / \partial x^i = 0 \qquad (\rho = 0, 1, ..., n). \tag{2.14}$$

Let K be the set of all vectors k_ρ of the form

$$k_\rho = F^\rho(t, x_0(t)) - F_0{}^\rho(t, x_0(t)) \qquad (\rho = 0, 1, ..., n), \tag{2.15}$$

where t is a point of continuity of $f_0(t, x_0(t))$ on (t^0, t^1) and F^ρ is given by (2.9) for $\rho = 0$ and by (2.12) for $\rho \geqslant 1$.

LEMMA 2.1. *The class K is a derived set for J^ρ at X_0.*

Suppose for a moment the lemma has been proved. Then, by Theorem 6.1 of [2], there exists multipliers $\lambda_0 \geqslant 0, \lambda_1, ..., \lambda_n$, not all zero, such that the inequality

$$L(k) = \lambda_\rho k^\rho \geqslant 0 \qquad (\rho = 0, 1, ..., n)$$

holds for every vector k in the closure of K.

Setting

$$F = \lambda_\rho F^\rho,$$

where F^ρ is obtained from (2.15), we see that

$$F(t, x_0(t)) \geqslant F_0(t, x_0(t)) \tag{2.16}$$

for all f in $\mathscr{F}$ except possibly at discontinuities of $f_0(t, x_0(t))$. By continuity considerations it holds there also.

Set

$$H(t, x, f, p) = p_i f^i - \lambda_0 f^0$$

where

$$p_i(t) = -\lambda_j P_{ij}(t) - \lambda_0 q_i(t).$$

Then (2.3) and (2.4) follow from the definitions of $P_{ij}(t)$, $q_i(t)$ and F^ρ and inequality (2.16).

We now return to the proof of Lemma 2.1. Let $k_1, ..., k_N$ be a set of N vectors of the form (2.15) in K. Then

$$k_j{}^\rho = F_j{}^\rho(t_j, x_0(t_j)) - F_0{}^\rho(t_j, x_0(t_j)) \tag{2.17}$$

where we can assume

$$t_1 \leqslant t_2 \leqslant \cdots \leqslant t_N.$$

Define

$$T_1 = t_1, \, T_j = t_j + \epsilon_1 + \cdots + \epsilon_{j-1} \qquad (j = 1, ..., N),$$

where $0 \leqslant \epsilon_i \leqslant \delta$. We choose δ so that $t_i + N\delta < t_1$ and $T_i < T_j$ if $t_i < t_j$. Let $M(\epsilon)$ be the complement on $[t^0, t^1]$ of this set of intervals.

Let $f_j(t, x)(j = 1,..., N)$ be the functions appearing in F_j in (2.17). Set

$$\begin{aligned}\phi^i(t, x, \epsilon) &= f_j^i(t, x) \qquad (T_j \leqslant t \leqslant T_j + \epsilon_j; j = 1,..., N)\\ &= f_0^i(t, x) \qquad (t \text{ in } M(\epsilon)).\end{aligned}$$

Then the equations

$$\dot{x}^i = \phi^i(t, x, \epsilon),\ x^i(t^0) = X^{i0}$$

have solutions

$$x^i(t, \epsilon) \qquad (t^0 \leqslant t \leqslant t^1; 0 \leqslant \epsilon \leqslant \delta)$$

for δ sufficiently small. Also $x(t, 0) = x_0(t)$, and the derivatives $x^i_{\epsilon_j}(t, \epsilon)$ are uniformly bounded piecewise continuous functions of t. The functions

$$\Psi_\rho(\epsilon) = J_\rho(x(\epsilon)) - J_\rho(x_0)$$

are of class C' for $\epsilon \leqslant \delta$. To prove the lemma we need only prove that at $\epsilon = 0$ the relation

$$\partial\Psi_\rho/\partial\epsilon_j = k_j^\rho \qquad (j = 1,..., N)$$

holds (see [2], Section 6). Setting $\epsilon_i = 0$ $(i \neq j)$, we see that

$$\Psi_\rho(\epsilon) = \Phi_\rho(\epsilon) + \theta_\rho(\epsilon)$$

where

$$\Phi_\rho(\epsilon) = \int_{t_j}^{t_j+\epsilon_j} [F_j^\rho(t, x(t, \epsilon)) - F_0^\rho(t, x_0(t))]\, dt$$

and

$$\theta_\rho(\epsilon) = \int_{M(\epsilon)} [F_0^\rho(t, x(t, \epsilon)) - F_0^\rho(t, x_0))]\, dt.$$

Note that $\Psi_\rho(0) = 0$. We have that at $\epsilon = 0$

$$\begin{aligned}\partial\Phi_\rho/\partial\epsilon_j &= \lim_{\epsilon_j=0} \Phi_\rho(\epsilon)/\epsilon_j = k_j^\rho,\\ \partial\theta/\partial\epsilon_j &= 0\end{aligned}$$

where the last equation follows from (2.12) and the boundedness of the partial derivative of $x(t, \epsilon)$ with respect to ϵ_j.

This completes the proof of Lemma 2.1, hence also of Theorem 2.1. In the special case where $\mathscr{F}$ is the set of functions given by

$$f(t, x) = g(t, x, u(t))$$

for $u(t)$ in an appropriate class of control functions, the conclusions of Theorem 2.1 reduce to the conditions for optimal control problems as given in [2].

References

1. T. Guinn *J. Computer Syst. Sci* **1**, 227-234 (1967)—preceding paper in this issue. Also, Paper Number 4103—Tech. Rept., Douglas Aircraft Company, Santa Monica, California.
2. M. R. Hestenes. On Variational Theory and Optimal Control Theory. *SIAM J. Control* **3**, 23-48 (1965).
3. T. Guinn. Weakened Hypotheses for the Variational Problem Considered by Hestenes. *SIAM J. Control.* **3**, 418-423 (1965). Also, Paper Number 3306—Tech. Rept., Douglas Aircraft Company, Santa Monica, California.